



Analyse d'une année d'enregistrements des températures à pas horaire pour établir un modèle spatial d'interpolation

Daniel Joly, Christian Debord, Marc Raynal

► To cite this version:

Daniel Joly, Christian Debord, Marc Raynal. Analyse d'une année d'enregistrements des températures à pas horaire pour établir un modèle spatial d'interpolation. XXVème colloque de l'Association Internationale de Climatologie: "Les climats régionaux : observation et modélisation", Sep 2012, Grenoble, France. pp.403-408. hal-00936785

HAL Id: hal-00936785

<https://hal.science/hal-00936785>

Submitted on 12 Feb 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ANALYSE D'UNE ANNÉE D'ENREGISTREMENTS DES TEMPÉRATURES À PAS HORAIRE POUR ÉTABLIR UN MODÈLE SPATIAL D'INTERPOLATION

Daniel JOLY ¹, Christian DEBORD ² et Marc RAYNAL ²

¹ ThéMA UMR 6049, CNRS et UFC, 32 rue Mégevand, 25030 Besançon cedex, France.

Tél. : 0381665402 ; daniel.joly@univ-fcomte.fr

² Institut Français de la Vigne et du Vin, Vinopôle, 39 rue M. Montaigne, 33290 Blanquefort, France.

Tél. : 05 56 35 58 80 ; Christian.DEBORD / Marc.RAYNAL@vignevin.com

Résumé : Une méthode d'analyse spatiale permettant de construire des champs continus est appliquée à la région viticole du Bordelais. La chronique des températures enregistrées à pas horaire de septembre 2010 à août 2011 en 78 stations sert d'abord de base à des régressions qui permettent d'identifier les variables significatives au seuil de 5 % parmi 13 variables spatiales. Le calcul de leur fréquence moyenne pour chaque heure de chaque mois indique que neuf d'entre elles sont très structurantes : distance à l'océan (52 %), altitude (37 %), pente (8 %) et, déclinée selon trois niveaux d'échelle, ampleur des creux (4 %) et des bosses (2,3 %). La fréquence et le gradient moyens sont ensuite utilisés pour calculer une température en tout point de grille, pour chacune des 24 h de chaque mois. La comparaison avec les températures observées montre que l'écart type d'erreur est de 0,5°C.

Mots-clés : Températures, variables explicatives, vignoble bordelais

Abstract: One year of hourly temperature records analysis to establish a spatial interpolation model.

A spatial analysis method for constructing continuous fields is applied to the vine region of Bordeaux. Temperature hourly recorded from September 2010 to August 2011 in 78 stations provides the basis for regressions that identify variables significant at the 5% threshold among 13 spatial variables. Calculating their average frequency for each hour of each month indicates that nine of them are very structuring: distance to the ocean (52%), altitude (37%), slope (8%) and, calculated for three scale levels, amplitude of hollows (4%) and humps (2.3%). The frequency and gradient means are then used to calculate a temperature at any grid point for each of the 24 hours of each month. The standard deviation of error is 0.5°C.

Keywords: Temperature, explanatory variables, Bordeaux vineyard.

Introduction

L'Institut Français de la Vigne et du Vin (IFV), centre technique interprofessionnel, a pour vocation de conduire des études de portée générale pour l'ensemble de la filière viti-vinicole. L'un des objectifs du groupe chargé des risques d'épidémie sur le vignoble est de modéliser les conditions propices au développement des principales maladies (mildiou, oïdium, black rot, ...) qui provoquent des destructions parfois très sévères du vignoble. Le projet Ecophyto vise ainsi à limiter la consommation de produits phytosanitaires. Les modèles mathématiques utilisés nécessitent que soient connues les températures à pas horaire de septembre à août pour estimer le risque et ainsi prévoir la survenue imminente de maladie. Ces modèles sont d'ordinaire appliqués aux postes climatiques. Ils peuvent l'être à l'ensemble du vignoble pour peu que les températures soient connues partout : c'est là que réside l'objectif prioritaire de cette étude qui porte sur la région viticole du Bordelais.

Classiquement, les interpolations reposent sur des techniques statistiques (régressions) ou géostatistiques (krigeage). La méthode d'interpolation, ici, est différente car elle s'appuie sur les résultats d'une modélisation spatiale préalable. L'idée est d'analyser la manière dont les variables topographiques structurent la variation spatio-temporelle des températures afin d'en déceler les régularités. Pour ce faire, une première étape portera sur des régressions pour identifier le gradient et la fréquence avec lesquels neuf variables explicatives expliquent significativement la température sur l'aire d'étude. Ces valeurs serviront ensuite à paramétrer les indices qui permettront d'estimer les températures. Des modèles horaires moyens seront

établis pour chacun des douze mois (septembre 2010 - août 2011) au cours desquels l'étude a été menée. Enfin, ce sont ces résultats qui seront appliqués en vue d'interpoler les températures en tout point de l'espace.

La première section concerne la nature des données et la méthode utilisée pour déterminer le gradient et la fréquence des variables qui expliquent significativement la variation des températures. La seconde section présentera les résultats en limitant le propos à la variation horaire et mensuelle des fréquences et du signe du gradient des cinq principales variables jouant un rôle significatif sur la variation spatiale des températures. La troisième section critiquera les modélisations spatiales obtenues en utilisant les résultats (fréquence et signe) de la section 2.

1. Données et méthode

L'analyse par régression des valeurs thermiques requiert deux types de données : les données collectées dans des stations climatiques et les données géographiques décrivant l'environnement de chaque station. Ces variables explicatives ont été choisies en fonction de nos connaissances préalables des processus qui règlent l'organisation spatiale de la température (Joly *et al.*, 2012). Les données de température concernent l'aire couverte par le vignoble dans les départements de Gironde et de Dordogne. Elles ont été enregistrées en 78 stations d'un réseau professionnel, DEMETER, géré par la « FREDON Aquitaine », et déployé sur le vignoble, du 1^{er} septembre 2010 au 31 août 2011. Au total, l'analyse a porté sur 8760 situations thermiques horaires en vue d'identifier les variables explicatives les plus fréquemment significatives. Elle a consisté à régresser les températures sur treize variables calculées à partir d'une image de l'occupation du sol et d'un MNT de l'IGN à pas de 50 m :

- la distance à la mer ;
- cinq variables topographiques (altitude, rugosité, pente, cosinus et sinus des versants) calculées dans la cadre d'une fenêtre circulaire dont le diamètre est de 3 pixels (Joly et Brossard, 2007) ;
- le rayonnement global (fenêtre 3*3) ;
- les formes topographiques sont décrites selon les deux modalités suivantes : les formes en relief, émergentes, que nous appellerons « bosses » et les formes en creux (Joly *et al.*, 2012). Leur ampleur est déclinée selon trois échelles : ont été identifiées les formes à grande échelle (250 m, fenêtre 5*5), à échelle moyenne (1750 m, fenêtre 35*35) et à petite échelle (5050 m, fenêtre 101*101).

L'intérêt de la démarche proposée est qu'elle procède d'une manière expérimentale pour établir le degré de significativité de ces variables et pour calibrer leur contribution à l'explication de variation spatiale de la température. Les résultats pourraient être faussés s'il existait des variables colinéaires. De fait, l'altitude covarie avec la distance à la mer ($r = 0,70$), ou avec l'ampleur des bosses ($r = 0,53$), le rayonnement global avec le cosinus des versants ($r = -0,81$). Nous avons résolu ces problèmes de multi colinéarité en recourant à la corrélation partielle. Une corrélation partielle est la corrélation entre deux variables lorsque l'influence de l'une sur l'autre est annulée. Il s'agit d'une statistique particulièrement utile lorsqu'on cherche à interpréter les relations entre plusieurs variables en termes de causalité (Foucart, 2006), ce qui est notre cas. Le fait que la colinéarité puisse changer le signe d'un coefficient est gênant pour étudier l'effet propre de toute variable, sur la température. Grâce à la corrélation partielle, chaque variable explicative est dégagée de l'influence que pourrait avoir chacune des douze autres variables sur elle et, ainsi, les résultats ne seront pas biaisés. Le coefficient de corrélation partielle permettra de juger la performance des modèles

statistiques ainsi que de calculer la fréquence avec laquelle chaque variable est significative au seuil de 5 %, heure par heure et pour chacun des 12 mois analysés, indépendamment les uns des autres.

2. Résultats

Nous avons réalisé 8760 analyses. Pour chacune d'entre elles, les treize variables explicatives ont été régressées indépendamment les unes des autres sur les températures observées aux 78 stations. Il en ressort treize coefficients de corrélation linéaire pour chacune des 8760 analyses. Les variables qui obtiennent un r significatif au seuil de 5 % sont incrémentées en gardant en mémoire l'heure et le mois auxquels elles appartiennent. A l'issue des analyses, la somme des occurrences est transformée en fréquence. La valeur du coefficient de régression (pente qui est le gradient altitudinal pour l'altitude par exemple) est également mémorisée quand la variable est significative. Une moyenne horaire pour chacun des douze mois est établie. Les résultats présentés ci-dessous concernent des moyennes annuelles. Globalement, de septembre 2010 à août 2011, deux variables se dégagent nettement : la distance à l'océan et l'altitude expliquent significativement la température dans respectivement 52 et 37 % des situations. Elles sont suivies par la pente (8 %), l'ampleur des creux (4 %) et des bosses (2,3 %). La fréquence de toutes les autres variables est inférieure à 2 %.

2.1. Distance à l'océan, altitude et pente

La fréquence avec laquelle la distance à la mer explique la température présente un mode (60 %) centré sur 14 h (figure 1A) et un minimum (40 %) au début de la matinée. Les gradients positifs très abondants durant le jour (50 %, soit près de 90 % des occurrences : pendant la journée, quand la distance à l'océan influence la température, cette dernière croît presque toujours à mesure que l'on s'enfonce à l'intérieur des terres). Inversement, durant la nuit, le gradient est très souvent négatif (la température est plus froide à l'intérieur des terres qu'au bord de l'océan). Ce comportement est conforme au schéma d'une terre qui s'échauffe, au cours de la journée, et se refroidit, au cours de la nuit, plus que l'océan et la frange côtière.

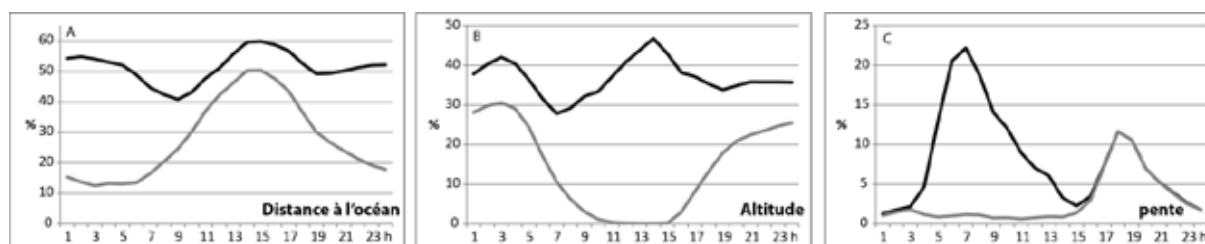


Figure 1 : Variation horaire de la fréquence avec laquelle la distance à la mer, l'altitude et la pente expliquent significativement la température ; en noir = fréquences totales, en gris : fréquence des gradients > 0.

L'influence de l'altitude sur les températures présente un schéma analogue avec un creux et un pic qui se produisent deux heures plus tôt (figure 1B). Le gradient altitudinal est toujours négatif au cœur de la journée (la température diminue avec l'altitude) et très souvent positif durant la nuit. Cette fréquence élevée d'inversions thermique pourrait être expliquée par les formes en creux (influence des vallées, quelle que soit leur échelle, où l'air froid stagne). Ce n'est pas le cas car l'influence de l'ampleur des creux sur l'altitude est contrôlée par la corrélation partielle. En l'état actuel de nos recherches, aucune hypothèse n'a été formulée pour expliquer ce phénomène qui rejoint des observations faites dans le même sens (Joly *et al.*, 2011 ; Joly, 2012).

Enfin, la pente (figure 1C) présente deux modes, l'un en début de matinée (23 %), le second (12 %) en fin d'après-midi. Durant la matinée, le signe de la corrélation est négatif (les pentes fortes sont plus froides que les pentes douces) tandis qu'en fin d'après-midi, il est exclusivement positif (plus la pente est vigoureuse, plus les températures sont élevées). Le rôle des vents adiabatiques et anabatiques est ainsi mis en évidence.

2.2. Ampleur des creux et des bosses

Concernant l'ampleur des creux et des bosses, quatre analyses ont été effectuées. Les trois premières portent sur chaque fenêtre, individuellement. La quatrième inclut les trois fenêtres en tant que variables explicatives de la température (l'analyse comporte alors 13 variables explicatives au lieu de neuf pour les trois premières analyses).

Globalement, l'ampleur des creux des deux plus grandes échelles (4,4 et 4,0 %) explique la température deux fois plus fréquemment que celle des bosses (1,9 et 1,8 %). La fréquence de l'ampleur des bosses à petite échelle augmente sensiblement sa fréquence (5,3 %) pour égaler presque celle des creux (6 %). Lorsque les trois fenêtres sont intégrées dans l'analyse, on ne comptabilise que l'échelle qui obtient le coefficient de corrélation le plus élevé. La fréquence des bosses et des creux est assez élevée (10 % pour les bosses et 14 % pour les creux). Ce score correspond, à peu de choses près, à la somme des fréquences individuelles obtenues par chacune des trois fenêtres. Cela signifie sans doute que les processus thermiques associés à ces trois échelles sont indépendants les uns des autres et qu'ils expliquent la température à tour de rôle.

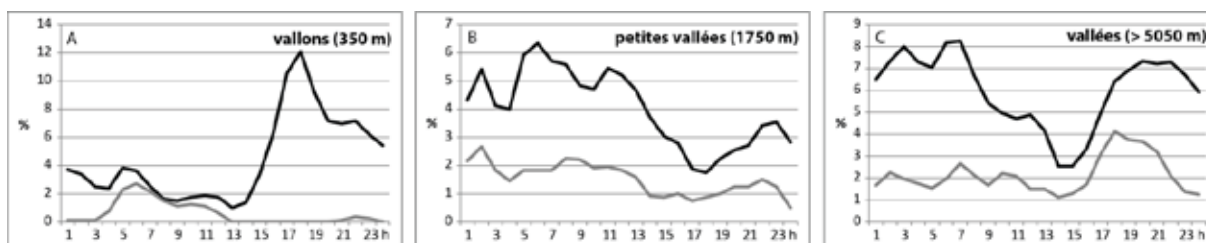


Figure 2 : Variation horaire de la fréquence avec laquelle de l'ampleur des creux explique significativement la température ; en noir = fréquences totales, en gris : fréquence des gradients > 0.

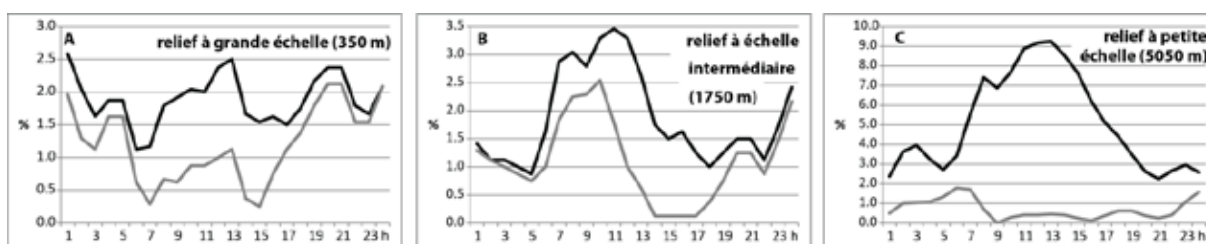


Figure 3 : Variation horaire de la fréquence avec laquelle de l'ampleur des bosses explique significativement la température ; en noir = fréquences totales, en gris : fréquence des gradients > 0.

La fréquence des creux à grande échelle en tant que variable explicative (figure 2A) est faible (1-4 %) jusqu'à 15 h et beaucoup plus élevée ensuite, avec un pic de 14 % à 18 h. Le signe de la corrélation est positif uniquement le matin. Les creux à plus petite échelle (figures 2B et 2C) expliquent plus fréquemment la température le matin et en soirée (avec un gradient le plus souvent négatif (plus les creux sont amples, plus la température est froide) qu'en fin d'après-midi ; le signe est alors positif dans environ 40 % des cas. L'opposition entre les creux à grande échelle (fréquence maximum en fin d'après-midi) et les creux aux deux autres échelles montre que la variation spatiale de la température dépend fortement de l'échelle des formes topographiques. La fin d'après-midi est un moment où la température est d'autant plus

faible que les vallons sont profonds alors que l'ampleur des vallées joue un rôle mineur sur les températures, comme si la turbulence thermique brassait plus facilement l'air des milieux topographiques ouverts que fermés. C'est l'inverse qui se produit durant la nuit, à un moment où l'air, plus calme s'accumule au fond des vallées en proportion de leur profondeur.

La fréquence des bosses oppose à nouveau les fréquences obtenues par la grande échelle (pas de variation cyclique ; figure 3A) aux échelles moyennes et petites (figures 3B et 3C) qui présentent une fréquence maximale centrée vers midi (3 %, 9 %) et des fréquences minimales durant la nuit. Le signe est dans presque tous les cas négatif le jour et positif durant la nuit.

A l'issue de ces analyses, neuf variables (distance à la mer ; altitude ; pente ; ampleur des bosses et des creux * trois niveaux d'échelle) sont identifiées comme exerçant une influence régulière sur la température. Pour chacune d'entre elles, nous avons calculé la fréquence et le gradient moyens pour chaque heure, en moyenne mensuelle. Il s'avère en effet que la valeur de l'une et l'autre de ces deux indices se modifie assez sensiblement selon les saisons. Cette variation n'est pas présentée ici faute de place.

3. Modélisation spatiale

Pour conduire la construction du champ continu de température par modélisation spatiale (et non par interpolation), nous disposons de 12 fichiers mensuels qui archivent chacun la fréquence et le gradient pour chaque heure du jour et chacune des neuf variables retenues. La température en chaque point de grille $p_{i,j}$ est d'abord déclarée égale à la moyenne thermique de l'heure et du mois de calcul. Ensuite, cette température est ajoutée au produit de la fréquence, du gradient et de la valeur de chaque variable en $p_{i,j}$. Pour une heure et un mois donnés, on aura, pour le point $p_{i,j}$ à une heure et un mois donnés :

$$\text{Temp}_{i,j} = f_1 * b_1 x_{1,i,j} + f_2 * b_2 x_{2,i,j} + \dots + f_9 * b_9 x_{9,i,j} + \overline{m}$$

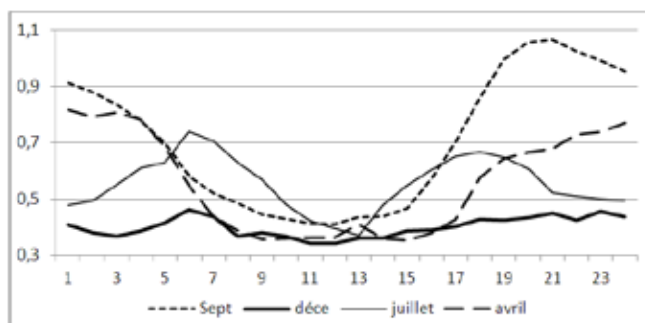
Où i, j = latitude, longitude de la grille spatiale ; f_1, \dots, f_9 = fréquence de la variable 1 à la variable 9 ; b_1, \dots, b_9 = gradient de la variable 1 à la variable 9 ; x_1, \dots, x_9 = valeur des variables 1 à 9 ; \overline{m} = moyenne de l'heure et du mois concernés.

La température obtenue sera inférieure à la moyenne si la somme des neuf produits fournit une valeur négative ; elle sera positive dans le cas contraire. Ainsi, une différenciation spatiale des températures prend place qui oppose les secteurs plus frais aux secteurs plus chauds que la moyenne. A titre d'exemple, nous décrirons les modèles spatiaux de 6 et 15 h en janvier et juillet 2011 (les cartes ne sont pas reproduites). Le fait saillant est le basculement ouest-est des pôles les plus chauds, du matin à l'après-midi, conformément à l'océan qui influence la température, notamment à 6 h en janvier (54 % des cas) et en juillet, à 15 h (64 %). Le gradient s'établit entre -0,027°C/km et +0,033°C/km. L'altitude joue un grand rôle également, surtout à 6 h en juillet en raison d'une fréquence élevée (29 %) et d'un gradient de -1,1°C/100 m. A l'exception de la pente à 6 h en juillet (fréquence de 48 % et gradient de -0,2°C/degré), l'influence des autres variables est plus discrète car leur fréquence est moindre.

Les points de grille qui supportent une station d'observation constituent un sous-ensemble des champs continus tels qu'ils viennent d'être réalisés. En effectuant la différence entre la température observée et la valeur du modèleau pixel qui supporte la station, on obtient une valeur de résidu qui permet de valider le résultat des modèles. L'écart-type des résidus, calculé sur la base des 78 stations et des 24 modèles horaires, est minimal en hiver (0,4°C) et maximal en été (0,6°C). Avec 0,4°C, ils sont plus faibles au cours de la journée (figure 4) et plus élevés durant la nuit (0,6°C en été, 0,8°C au printemps et 1°C en septembre). Il est probable que ces différences sont dues au brassage de l'air qui, en journée, estompe les différences locales de température. En hiver, les brouillards et situations perturbées

contribuent également à effacer les écarts thermiques. La conjonction de ces deux facteurs facilite l'interpolation.

Figure 4 : Variation de l'écart-type d'erreur pour les 24 modèles horaires et quatre mois.



Conclusion

Les régularités discernées dans la chronique des températures à pas horaire enregistrée pendant un an a permis d'identifier les variables qui expliquent le plus fréquemment la température. Il s'agit de la distance à la mer, de l'altitude et de la pente auxquelles il convient d'ajouter l'ampleur des bosses et des creux à trois niveaux d'échelle. La fréquence et le gradient moyens de ces variables ont été calculés pour chacune des 24 heures et pour chacun des 12 mois de la chronique. Ces deux valeurs ont permis de construire le champ continu de la température sur l'ensemble de l'aire couverte par le vignoble bordelais sur les départements de Gironde et de Dordogne. La comparaison avec les données enregistrées aux 78 stations a montré que les modèles spatiaux reconstituent bien les températures : l'écart type d'erreur est de 0,5°C seulement. De plus, il est probable que fréquences et gradients dépendent étroitement des conditions atmosphériques (la fréquence de la distance à la mer, ou de tout autre variable, en tant que variable explicative de la température n'est sans doute pas la même par temps anticyclonique que par temps perturbé d'ouest). Or nous avons travaillé en moyenne mensuelle, ce qui a sans doute contribué à mélanger plusieurs structures différentes et donc, à édulcorer, voire, à fausser le signal. Il serait du plus haut intérêt de segmenter l'information selon les types de temps les plus représentatifs et reproduire, pour chacun d'eux, les calculs qui ont été présentés ici. Le gain de précision attendu d'une telle expérience mériterait d'être précisé par des recherches complémentaires prenant en compte plusieurs années, ce qui permettrait, en outre, de vérifier la robustesse des résultats.

Pour autant, les valeurs produites par ces modèles ne vaudront jamais les estimations fournies par la méthode statistique (régression puis krigeage des résidus) car, en ce cas, le résultat des analyses est ajusté à chaque situation horaire. Mais s'il s'agit d'interpoler des températures concernant des prévisions de température (utilisant les sorties du modèle Safran par exemple), la méthode statistique n'est plus adaptée. Alors, notre méthode devient une alternative intéressante, la résolution spatiale de Safran étant trop grossière pour les applications agroclimatologiques envisagées.

Références bibliographiques

- Foucart T., 2006 : Colinéarité et régression linéaire. *Math. & Sci. Humaines*, 173(1), 5-25.
- Joly D., Brossard T., 2007 : Contribution of environmental factors to temperature distribution at different resolution levels on the forefield of the Loven Glaciers (Svalbard). *Polar Record*, 43(4), 353-359.
- Joly D., Bois B., Zaksek K., 2012 : Rank-ordering of topographic variables correlated with temperature. *Atmospheric and Climate Science*, 2(2), 139-147. DOI: 10.4236/acs.2012.22015DOI.
- Joly D., Brossard T., Cardot H., Cavailhès J., Hilal M., Wavresky P., 2011 : Temperature Interpolation by local information; the example of France. *International Journal of Climatology*, 31(14), 2141-2153.